

What is claimed is:

1. A method for identifying genes expressed during differentiation of a cell comprising the steps of:

- 5 a) integrating into a site in the genome of a host cell, a cell lineage targeting vector comprising, a pair of recombinase recognition sites flanking one or more polyadenylation sites, a first selectable marker placed downstream of or between the two recombinase recognition sites, a reporter gene placed downstream of the recombinase recognition sites, and a cell lineage specific gene promoter placed  
10 upstream of the recombinase recognition sites or a cell specific lineage gene placed downstream of the recombinase recognition sites,
- b) amplifying cells generated from the host cell;
- c) integrating into the genome of a plurality of the amplified cells, a gene-trap vector comprising a splice acceptor, a type IIS restriction endonuclease  
15 cleavage site, a recombinase, one or more polyadenylation sites, a second selectable marker and a splice donor;
- d) allowing the cells to differentiate;
- e) isolating cells in which the reporter gene is expressed indicating expression of the cell lineage specific gene;
- 20 f) identifying trapped genes in the isolated cells.

2. The method of claim 1, wherein the identification of trapped genes in the isolated cells comprises the steps of

- 25 a) preparing from isolated cells in 1e), concatamers comprising portions corresponding to trapped genes in the isolated cells;
- b) sequencing the concatamers to identify trapped genes;
- wherein the each trapped gene is indicative of a gene expressed during differentiation.

30 3. The method of claim 2, wherein the portions of trapped genes are amplified by inverse PCR.

4. The method of claim 2, wherein the portions of trapped genes are amplified by RT PCR.
5. The method of claim 1, wherein the step of identifying the trapped genes in  
5 step f) comprises the steps of:
- a) preparing mRNA from cells in which the fluorescent reporter is expressed in d);
  - b) synthesizing a first and second cDNA strands from the mRNA;
  - c) digesting with type IIS restriction endonucleases to produce Assay Tags
- 10 wherein each Assay Tag comprises a portion of a trapped gene and a portion of the gene-trap vector;
- d) concatenating the Assay Tags;
  - e) amplifying and sequencing the concatamers to identify the sequence of the portion of the trapped gene.
- 15
6. The method of claim 5, wherein the second DNA strand is biotinylated.
7. The method of claim 1, wherein the Type IIS restriction endonuclease is selected from the group consisting of BsgI, BpmI, BsmF1, MmeI and FokI.
- 20
8. The method of claim 1, wherein the reporter protein is a fluorescent protein.
9. The method of claim 8, wherein the fluorescent reporter protein is EGFP.
- 25
10. The method of claim 1, wherein the recombinase is Cre or FLP.
11. The method of claim 10, wherein the recombinase is fused to thymidine kinase or nitroreductase.
- 30
12. A method for identifying genes expressed during differentiation of a cell comprising the steps of:
- a) integrating into a site in the genome of a host cell, a cell lineage targeting

vector comprising a pair of recombinase recognition sites flanking one or more polyadenylation sites, a first selectable marker, a reporter gene, and a cell lineage specific gene promoter or a cell lineage specific gene, wherein recombinase based excision allows the expression of the reporter gene;

- 5           b) amplifying cells generated from the host cell;
- c) integrating into a plurality of the amplified cells, a gene-trap vector comprising a splice acceptor, a type IIS restriction endonuclease cleavage site, a recombinase, a second selectable marker, and either a splice donor or a polyadenylation site, wherein integration of the gene-trap vector into an endogenous
- 10       gene allows the recombinase to be produced and also incorporates a type IIS endonuclease site into the endogenous gene.
- c) allowing the host cells to differentiate;
- d) isolating cells in which the reporter gene is expressed indicating expression of the cell lineage specific gene;
- 15       e) digesting DNA from the isolated cells to form fragments comprising portions of trapped genes;
- f) concatenating and sequencing the fragments comprising portions of trapped genes.

20       13.     The method of claim 12, wherein the fragments of DNA comprising portions of trapped genes are amplified by inverse PCR.

14.     The method of claim 12, wherein the fragments of DNA comprising portions of trapped genes are amplified by RT PCR.

25       15.     A method for identifying genes expressed during differentiation of a cell comprising the steps of:

- a) integrating into a site in the host cell of a genome, a cell lineage targeting vector comprising a pair of recombinase recognition sites flanking one or more
- 30       polyadenylation sites, a first selectable marker, a reporter gene, and a cell lineage specific gene promoter or a cell lineage specific gene, wherein recombinase based excision allows the expression of the reporter gene;

- b) amplifying cells generated from the host cell in a)
- c) integrating into a plurality of the amplified cells, a gene-trap vector comprising a splice acceptor, a type IIS restriction endonuclease cleavage site, a recombinase, a second selectable marker, and either a splice donor or a polyadenylation site, wherein integration of the gene-trap vector into an endogenous gene allows the recombinase to be produced and also incorporates a type IIS endonuclease site into the endogenous gene.
- c) allowing the host cells to differentiate;
- d) isolating cells in which the reporter gene is expressed indicating expression of the cell lineage specific gene;
- e) preparing mRNA from cells in which the reporter gene is expressed in d);
- f) synthesizing a first and second cDNA strands from the mRNA;
- g) digesting with type IIS restriction endonucleases to produce Assay Tags wherein each Assay Tag comprises a portion of a trapped gene and a portion of the gene-trap vector;
- h) concatenating the Assay Tags;
- i) amplifying and sequencing the concatamers to identify the sequence of the portion of the trapped gene.
16. The method of claim 15, wherein the second DNA strand is biotinylated.
17. The method of claim 15, wherein the Type IIS restriction endonuclease is selected from the group consisting of BsgI, BpmI, BsmF1, MmeI and FokI.
18. The method of claim 15, wherein the reporter protein is a fluorescent protein.
19. The method of claim 18, wherein the fluorescent reporter protein is EGFP.
20. The method of claim 15, wherein the recombinase is Cre or FLP.
21. The method of claim 20, wherein the recombinase is fused to thymidine kinase or nitroreductase.